

Investigación y desarrollo de aplicaciones de Realidad Aumentada y Visión por Computador utilizando un framework multipropósito

Alejandro Mitaritonna^{1,2}, Juan Lestani¹, Silvana Olmedo¹, Leonardo Inza¹

¹ Laboratorio de Procesamiento de Imágenes y Visión Artificial
Universidad de Belgrano, Facultad de Ingeniería y Tecnología Informática
{alejandro.mitaritonna, juan.lestani, silvana.olmedo, leonardo.inza}@comunidad.ub.edu.ar

² Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF)
amitaritonna@citedef.gob.ar

RESUMEN

En este artículo se presenta el proyecto de investigación y desarrollo de un framework de software multipropósito de Realidad Aumentada (RA) y Visión Artificial (VA) que se está llevando a cabo en el Laboratorio de Procesamiento de Imágenes y Visión Artificial de la Facultad de Ingeniería y Tecnología Informática de la Universidad de Belgrano. Dicho framework servirá como base para la construcción de aplicaciones móviles y de escritorio para ser utilizadas en diferentes disciplinas tales como medicina, entretenimiento, industria, turismo y educación, entre las principales. En el presente artículo se detallan los trabajos realizados hasta el momento, los logros alcanzados y las nuevas líneas de investigación que han surgido a medida que se desarrolla el proyecto.

Palabras claves: *realidad aumentada, visión por computador, framework*

CONTEXTO

El diseño y desarrollo del framework surge como iniciativa de la FITI-UB con el fin de crear un laboratorio de Procesamiento de Imágenes y

Visión Artificial dedicado a la investigación y desarrollo de proyectos relacionados con la temática. Dentro de esta línea se ha desarrollado una tesis doctoral en el Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF) en el marco del Programa de Investigación y Desarrollo para la Defensa (PIDDEF) (aprobado con el Nro. 22/12), elaborado por la Subsecretaría de Investigación Científica y Desarrollo Tecnológico del Ministerio de Defensa. Asimismo se está elaborando una tesis de grado de una de las integrantes del equipo de investigación.

INTRODUCCION

El proyecto está enmarcado en el área de procesamiento de imágenes y visión artificial [Mit17]. En [Mit18] se describe la línea de investigación presentada originalmente; las líneas/grupos de trabajo formados a fin de abordar las temáticas centrales: estimación de posición y orientación (pose) y registración de objetos, reconocimiento y clasificación de objetos y motores de renderizado de objetos; y los conceptos fundamentales de RA y VA.

El trabajo de la definición y construcción de un framework que facilite el desarrollo de

aplicaciones de RA comenzó por identificar y definir las partes componentes de un sistema de este tipo. De acuerdo a [Aba11] y basándose en lo que propone [Azu97], una aplicación de RA mínimamente está compuesto por las siguientes partes (figura 1):

- *Captura de la escena real*: El video capturado puede utilizarse para tracking basado en visión, es decir basado en el análisis de la imagen mediante algoritmos de visión
- *Tracking del usuario*: puede realizarse mediante dispositivos específicos o puede realizarse tracking basado en visión para lo cual es necesaria la captura de la escena real.
- *Generación de la escena virtual*: se tiene un mundo virtual, con la información de la posición y orientación del participante se genera una vista acorde del mismo
- *Rendering*: se combinan las imágenes del mundo real con los objetos virtuales. Los objetos virtuales se renderizan y se proyectan en el dispositivo móvil.

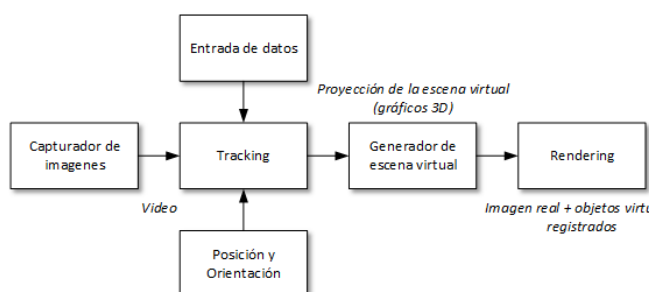


Figura 1

Adicionalmente se profundizó sobre los aspectos de VA, para ello se analizaron los componentes que forman parte del framework. Una de las particularidades del framework es que contará con un módulo de VA y cuya funcionalidad se centrará básicamente en el proceso de entender, interpretar y definir información visual de escenas en 2D y en 3D. Dicho módulo tendrá las siguientes características, entre las principales:

- Reconocimiento de Objetos

- Tracking de objetos basadas en características naturales
- Interfaz avanzada basada en gestos

Para concretar esta funcionalidad se está investigando y desarrollando pruebas de concepto cuyo objetivo es entender en profundidad la utilización de tecnologías de aprendizaje profundo para la clasificación de objetos. A continuación se detallan las principales líneas de investigación y desarrollo que el grupo de investigadores están realizando en el Laboratorio de Procesamiento de Imágenes y Visión Artificial de la Facultad de Ingeniería y Tecnología Informática de la Universidad de Belgrano.

LINEAS DE INVESTIGACIÓN Y DESARROLLO

Modelo de referencia

Como primera aproximación al desarrollo de un framework se planteó que su arquitectura respondiera a las exigencias de un modelo de referencia para aplicaciones de RA [Mit04a]. Un modelo de referencia en ingeniería de sistemas y en ingeniería de software es un modelo que contiene un objetivo o idea básica de algo, y que se puede establecer como una referencia para múltiples propósitos. Un modelo de referencia consiste de un conjunto mínimo de conceptos, axiomas y relaciones propios de un dominio particular de problema, y es independiente de estándares específicos, tecnologías, implementaciones, o de cualquier otro detalle concreto. Un estudio de varios sistemas de RA reveló que a pesar de ser bastante diferentes en los detalles, la mayoría de los sistemas de RA existentes comparten una estructura arquitectónica en común. El modelo de referencia se organiza en seis subsistemas lógicos:

- *Subsistema de seguimiento (tracking)*: Responsable de responder a los cambios de posición y orientación del usuario.
- *Aplicación*: Responsable del control de flujo, la lógica de la aplicación y la coordinación.
- *Modelo del mundo*: provee acceso a la representación digital del mundo.
- *Presentación*: Sistema responsable de la presentación de los resultados. Incluye renderizado 2D/3D y salida de audio o táctil.
- *Contexto*: provee a la aplicación con contexto tanto del estado como de la situación del usuario.
- *Interacción*: representa la interacción del usuario con el sistema de RA.

De acuerdo a lo antes mencionado, en la figura 2 se muestra la vista de componentes del modelo referencial propuesto por [Mac05]:

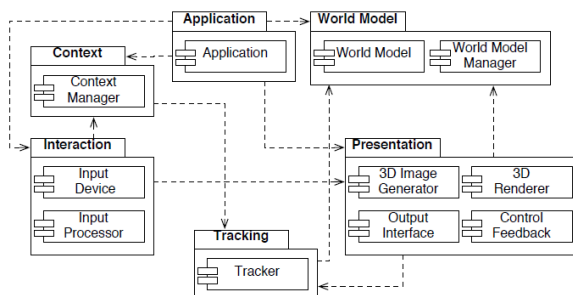


Figura 2

Cada subsistema está formando por varios componentes. El modelo de referencia describe a nivel abstracto los componentes generales y la estructura de los sistemas de realidad aumentada. Sin embargo, dependiendo de los requisitos funcionales de un sistema particular, algunos de los componentes pueden dejarse afuera [Mac04].

Pruebas de concepto y desarrollo de subsistemas

Con el objetivo de ir entendiendo y experimentando con la librería OpenCV, se realizaron pruebas de concepto, entendidas

como subproyectos del proyecto principal ya que sus resultados aportaban componentes y conocimientos para el desarrollo del framework. A continuación se describen brevemente las más importantes:

Reconocimiento de Marcadores

Una de las técnicas más difundidas para la determinación de la pose (y en general para situar objetos 3D en el mundo real) es la que está basada en el reconocimiento de marcadores. Éstas son figuras que contienen rasgos geométricos fácilmente detectables (bordes, esquinas, etc.) que, en su forma canónica, ya están almacenadas en el sistema, por lo que a partir de detectar su posición en la escena real, se pueden calcular las variaciones respecto de la forma almacenada y las transformaciones geométricas necesarias para representar esta variación o lo que es lo mismo, la relación entre el sistema de coordenadas centrado en la cámara y el del mundo real. De esta manera se determina la pose de la cámara.

Los pasos involucrados en el reconocimiento de marcas son complejos y abarcan [Mit14b]:

- Captura y conversión del frame de entrada
- Transformación a escala de grises
- Binarización según un umbral
- Detección del contorno
- Determinación de bordes y vértices
- Cálculo de las coordenadas de la cámara relativas a la marca

Sistema de ayuda de navegación y obtención de información de aulas en entornos delimitados

La aplicación se concibió como un sistema de ayuda para la navegación y obtención de información académica dentro del edificio de la Facultad de Tecnología Informática. Se colocaron diferentes códigos QR para cada aula y laboratorio de la facultad y se desarrolló una

aplicación móvil para smartphones funcionando bajo sistema Android.

En la primera etapa se implementó una aplicación que permitía identificar los espacios de trabajo enfocando con el smartphone el frente del aula o laboratorio de forma que el lector de códigos QR pudiera recuperar la información codificada en el marcador y presentarla en la pantalla del dispositivo como información de texto. Dicha información que se presentaba estaba formada por: información de laboratorios/aulas, asignaturas que se dictan en los mismos, horarios y profesores a cargo (figura 3).



Figura 3

La última etapa, aún en desarrollo, permitiría poner un objetivo espacial que se quiere alcanzar (aula o laboratorio), y la lectura de los marcadores que el navegante fuera encontrando en el entorno delimitado permitiría generar y superponer a la imagen captada por el dispositivo información de dirección a tomar (flechas) para llegar al destino.

Reconocimiento de objetos utilizando Deep Learning

El subproyecto, actualmente en curso, consiste en el desarrollo de un sistema de reconocimiento

de objetos reciclables basado en Redes Neuronales Convolucionales [Ant17]. Este tipo de redes son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones. Al momento de redactarse este artículo se está trabajando en la definición de algunas características de la arquitectura y se están reuniendo datos gráficos para armar una base de imágenes para el entrenamiento de la red. La función de la red será la de clasificar un objeto doméstico en alguna de 6 categorías objetos (p.ej.: envases). Este proyecto está sustentado por el trabajo de Tesina Final que está desarrollando una alumna del grupo de investigación del Laboratorio.

LINEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Diseño y desarrollo de un framework de RA y VA multipropósito (aplicado a diferentes disciplinas)
- Desarrollo de aplicaciones de RA y VA utilizando el framework propuesto
- Investigación de técnicas de Deep Learning aplicables al procesamiento de imágenes y clasificación de objetos utilizando como soporte el framework Keras [Ant17][Ker18]

RESULTADOS Y OBJETIVOS

A continuación se detallan los resultados obtenidos y los objetivos generales del proyecto:

- Se ha conformado un grupo de investigación y un Laboratorio dedicados al Procesamiento de Imágenes y Visión Artificial de la FITI en UB.
- El proyecto principal apunta a la creación de herramientas integradas en un entorno informático (framework) que facilite el desarrollo de aplicaciones de RA y VA.
- Dentro de este proyecto marco se estudian y exploran distintas tecnologías y

herramientas para entrenamiento de los sistemas en el reconocimiento de imágenes, generación de objetos. Estas pruebas de concepto fueron desarrolladas con el fin de ir materializando los conocimientos teóricos que se iban adquiriendo.

- Consistente con la potencialidad de aplicaciones basadas en el uso de Redes Neuronales Convolucionales y la disponibilidad de bases de imágenes rotuladas, aptas para el entrenamiento de estas redes mediante Deep Learning, se ha decidido dedicar una parte importante del proyecto a profundizar en el área de Reconocimiento de Objetos y desarrollar arquitecturas de redes multicapa capaces de clasificar una imagen entre un número limitado de categorías de objetos.
- Curso de Posgrado en Interfaces Avanzadas: Del mundo Real al Virtual. Departamento de Posgrado y Educación Continua. Facultad de Ingeniería y Tecnología Informática. Universidad de Belgrano.
- En el mes de noviembre del 2018 se presentó en el Instituto Inmaculada Concepción las capacidades de esta tecnología y los logros alcanzados en el laboratorio.
- Capacitación en el uso de procesamiento de imágenes, visión artificial y aprendizaje profundo a la empresa CyS en el marco del proyecto denominado Mapas Mentales financiado por el FONTAR.

FORMACIÓN DE RECURSOS HUMANOS

Uno de los principales objetivos de este proyecto de investigación y de la creación del Laboratorio de Procesamiento de imágenes y Visión Artificial es la formación de recursos en las temáticas de Realidad Aumentada y Visión Artificial enfocándose particularmente en Redes Neuronales Convolucionales para la clasificación de imágenes.

Durante el período de trabajo se terminó de elaborar una tesis de doctorado y se avanzó en la redacción de una tesina de grado:

- Alejandro Mitaritonna. “Realidad Aumentada para la Identificación de Objetivos Militares” (culminación de tesis de doctorado en la UNLP. Director: Dra. María José Abásolo; Codirector: Dr. Francisco Montero Simarro)
- Silvana Olmedo. “Sistema de Reconocimiento de objetos reciclables” (tesina de grado en Ingeniería Informática en la FITI-UB. Tutor: Mag. Alejandro Mitaritonna)

Asimismo se han capacitado a los alumnos que participan en el proyecto en el marco de esta línea de investigación: Carolina Páez, Florencia Vela, Tomás Poeta, Martín Lorenzo y Leonardo Inza.

Por último, nuestra intención durante el presente año es realizar demostraciones y presentaciones en Colegios u otros establecimientos de estudios e inclusive de investigación del uso de esta tecnología y los avances logrados en el Laboratorio.

REFERENCIAS

[Aba11] M. J. Abásolo Guerrero, C. Manresa Yee, R. Más Sansó y M. Vénere, Realidad virtual y realidad aumentada. Interfaces avanzadas., La Plata, Buenos Aires, Argentina: Editorial de la Universidad Nacional de La Plata (Edulp). 47 N° 380 / La Plata B1900AJP / Buenos Aires, Argentina, 2011.

[Ant17] Antona Cortés, C. Herramientas modernas en redes neuronales: La librería Keras. Trabajo Final de Grado. Univ. Autónoma de Madrid, 2017

[Azu97] R. T. Azuma. A Survey of Augmented Reality. *Teleoperators and Virtual Environments* 6, 4, 355-385, 1997.

[Ker18] Keras: The Python Deep Learning library. <https://keras.io/>

[Mac04] A. MacWilliams, T. Reicher, G. Klinker y B. Bruegge, *Design Patterns for Augmented Reality Systems*, Institut für Informatik, Technische Universität München: In Proc. International Workshop Exploring the Design and Engineering of Mixed Reality Systems - MIXER 2004, 2004.

[Mac05] A. MacWilliams, *A Decentralized Adaptive Architecture for Ubiquitous Augmented Reality System*, Institut für Informatik der Technischen Universität München, 2005.

[Mcd03] C. McDonald, *Hand Interaction in Augmented Reality*, Ottawa, Ontario, Canada: The Ottawa-Carleton Institute for Computer Science. School of Computer Science. Carleton University, 2003

[Mil94] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays", *IEICE Transactions on Information Systems*, Vol E77-D, Nr 12, Dec 1994.

[Mit14a] Mitaritonna, A., Pandolfo, L., Yokhdar D., Esteves, C. "RAIOM. Introducción a la arquitectura del framework de Realidad Aumentada". Escuela y Workshop Argentino en Ciencias de las Imágenes. ECIMAG 2014. Facultad de Ingeniería, UBA, 11 al 15 Agosto 2014.

[Mit14b] Mitaritonna, A., Pandolfo, L., Yokhdar D., Esteves, C. "RAIOM. Introducción a los algoritmos de visión por computador". Escuela y Workshop Argentino en Ciencias de las

Imágenes. ECIMAG 2014. Facultad de Ingeniería, UBA, 11 al 15 Agosto 2014.

[Mit15] Mitaritonna, A.; Abásolo Guerrero, M. J. (2015) *Improving Situational Awareness in Military Operations using Augmented Reality*. Proceedings of WSCG 2015. ISBN N°:978-80-86943-72-5, 2015

[Mit17] Mitaritonna, A.; Lestani, Juan. *Framework multipropósito de Realidad Aumentada y de Visión Artificial*. Congreso TE&ET, Universidad Nacional de La Matanza, Buenos Aires, Argentina, 2017.

[Mit18] Mitaritonna, A.; Lestani, Juan; Tarulla, Francisco; Poeta, Tomás; Olmedo, Silvana; Paez, Carolina; Lorenzo, Martín; Vela, Florencia; Inza, Leonardo. "Realidad Aumentada y Visión Artificial. Framework multipropósito". 20° WICC, UNNE, Argentina, 2018.